



LIRMM



Centre National de la
Recherche Scientifique



Université
Montpellier II

Analyse sémantique de textes et algorithmes à fourmis

LIRMM-INFO
Equipe TALN

Sélection/pondération d'acceptations (WSD)

“L’avocat plaide”

avocat/fruit ou avocat/justice ?

Rattachement de groupes prépositionnels

“Il voit la fille avec un télescope”

“Il (voit avec un télescope) la fille” ou

“Il voit la (fille avec un télescope)” ?

Interprétations multiples

“L’avocat est véreux” 2 mais pas 4 appariements ?

Résolution anaphorique

“L’avocat défend son client. Il sera acquitté”

il = avocat ou il = client ?

Instanciation des fonctions lexicales [meltchuk] [schwab]

“Il a une forte fièvre”

Magn(fièvre) = forte ?

Indexation de textes en RI [jaillet, prince, chauché, teissere]

Résolution de la polysémie lexicale (souris)
augmentation de la précision

Synonymie (chat/matou) / champs sémantique (cheval/équitation)
augmentation du rappel

Traduction Automatique [prince, delorme]

Résolution anaphorique - référent (he/she/it ? his/her/its ?)

Phénomènes contrastifs (river = rivière/fleuve ?)
(abats = ofals/giblets ?)

Fonctions lexicales (forte fièvre = high fever)

Analyse thématique et fourmis

Vecteurs conceptuels

Arbres morphosyntaxiques

Recherche de ressources et création de ponts

Couplage d'un réseau lexical

Relations ontologiques - relations prédicatives - typicalité

Castes de fourmis - agent / patient / ...

Pistes de recherche

Production entre castes - Auto-arrêt du système - Inhibition

Représentation thématique [chauché, lafourcade]

Item lexical = Idées = Vecteur conceptuel

Par exemple, 873 composantes (concepts issus du thésaurus Larousse)

(1)existence, (2)inexistence, (3)matérialité, ..., (516)liberté, ...,
(872)jeux, (873)jouets

Une composante du vecteur correspond à l'activation d'un concept.

Combinaison de vecteurs : addition, contextualisation...
[lafourcade, prince, schwab]

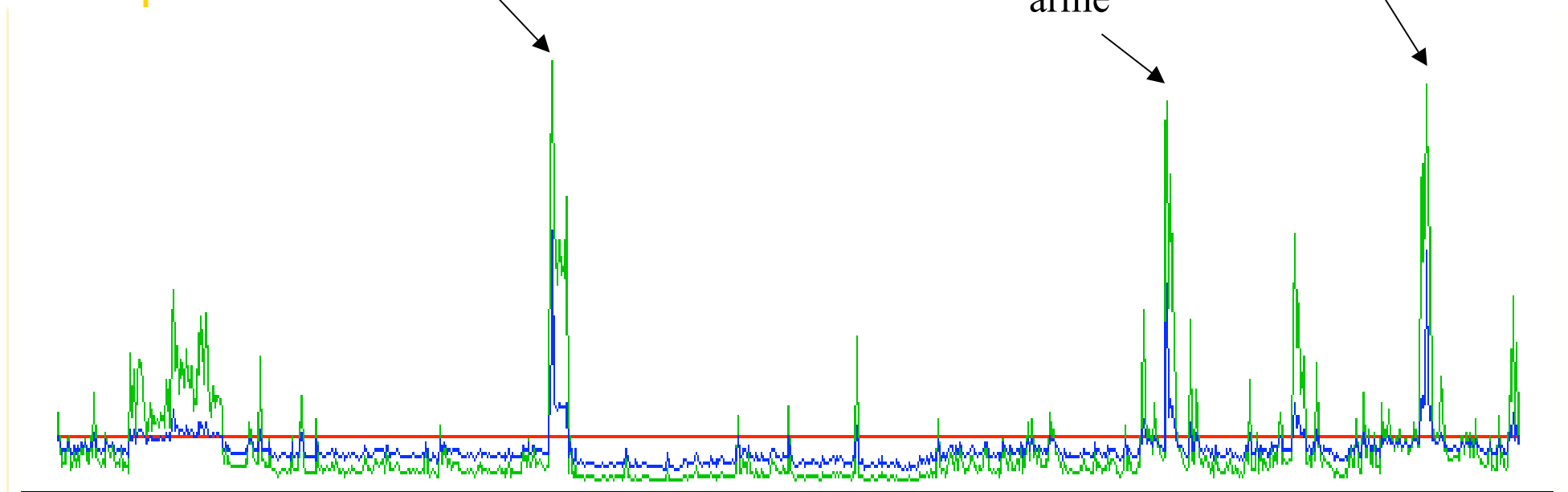
Vecteurs
conceptuels

frégate (3 sens fusionnés)

oiseau

arme

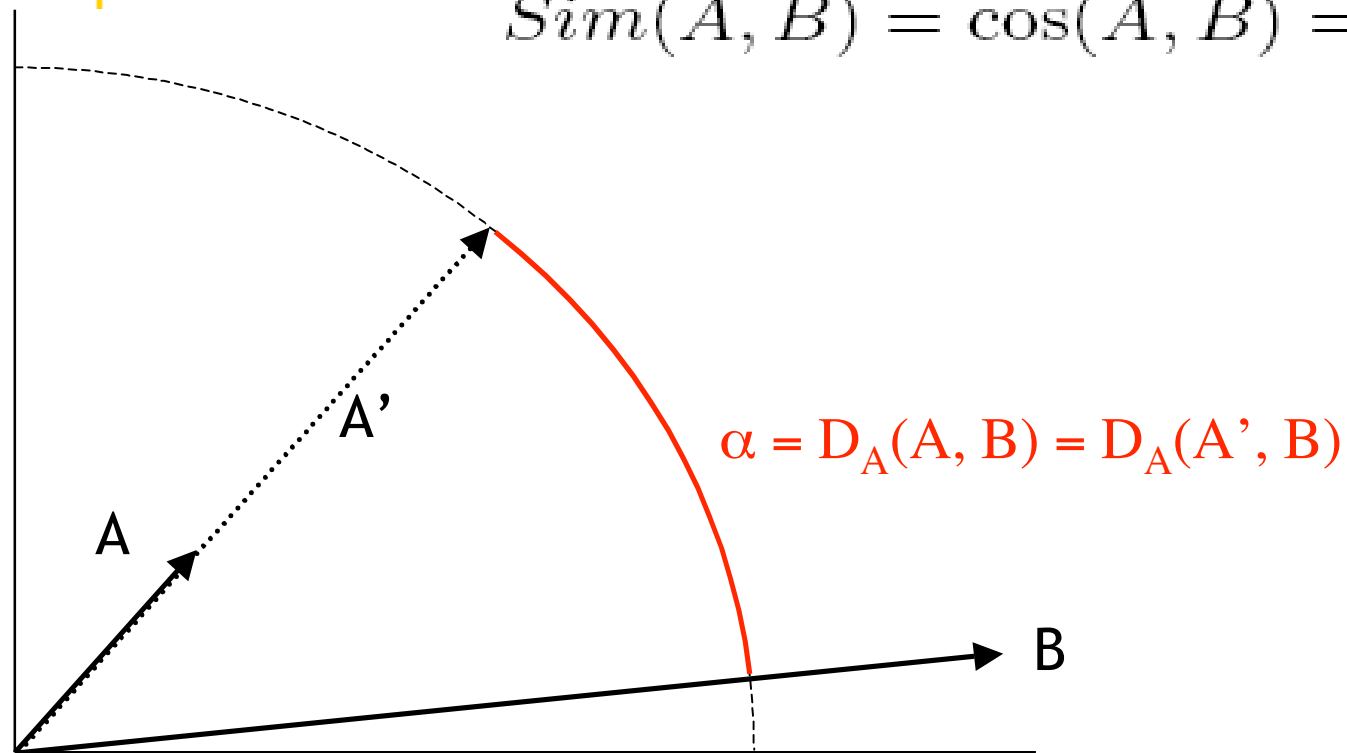
transports maritimes
et fluviaux



Vecteurs
conceptuels

$$D_A(A, B) = \arccos(\text{Sim}(A, B))$$

$$\text{Sim}(A, B) = \cos(\widehat{A, B}) = \frac{A \cdot B}{\|A\| \times \|B\|}$$



Vecteurs
conceptuels

$$D_A(\text{pélican}, \text{pélican}) = 0 \text{ (} 0^\circ \text{)}$$



$$D_A(\text{pélican}, \text{grand-gosier}) = 0,2 \text{ (} 11^\circ \text{)}$$



$$D_A(\text{pélican}, \text{train}) = 1,22 \text{ (} 70^\circ \text{)}$$



$$D_A(\text{pélican}, \text{oiseau}) = 0,46 \text{ (} 26^\circ \text{)}$$



$$D_A(\text{pélican}, \text{mouette}) = 0,4 \text{ (} 23^\circ \text{)}$$



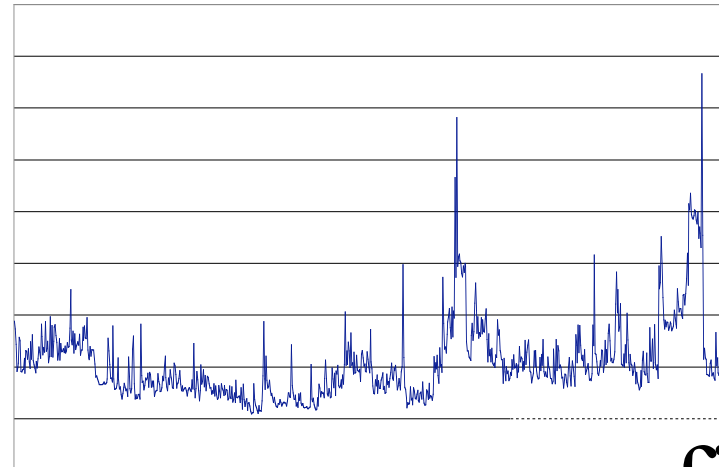
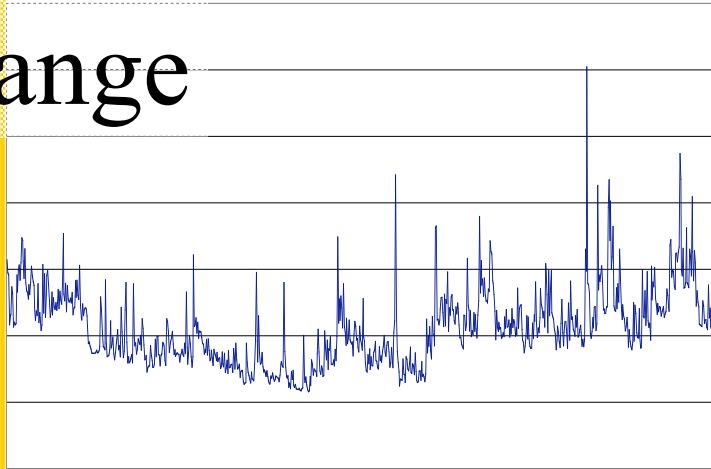
$$D_A(\text{pélican}, \text{poisson}) = 0,35 \text{ (} 20^\circ \text{)}$$



distance thématique \neq distance ontologique (de type *est-un*)
 mais distance thématique \supset distance ontologique

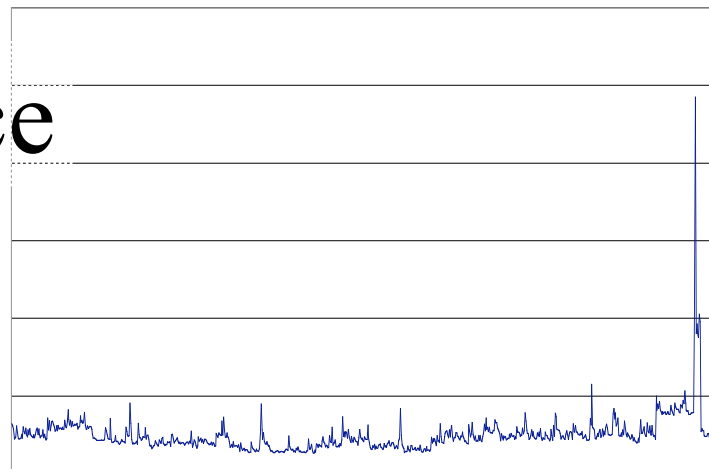


échange



profit

finance

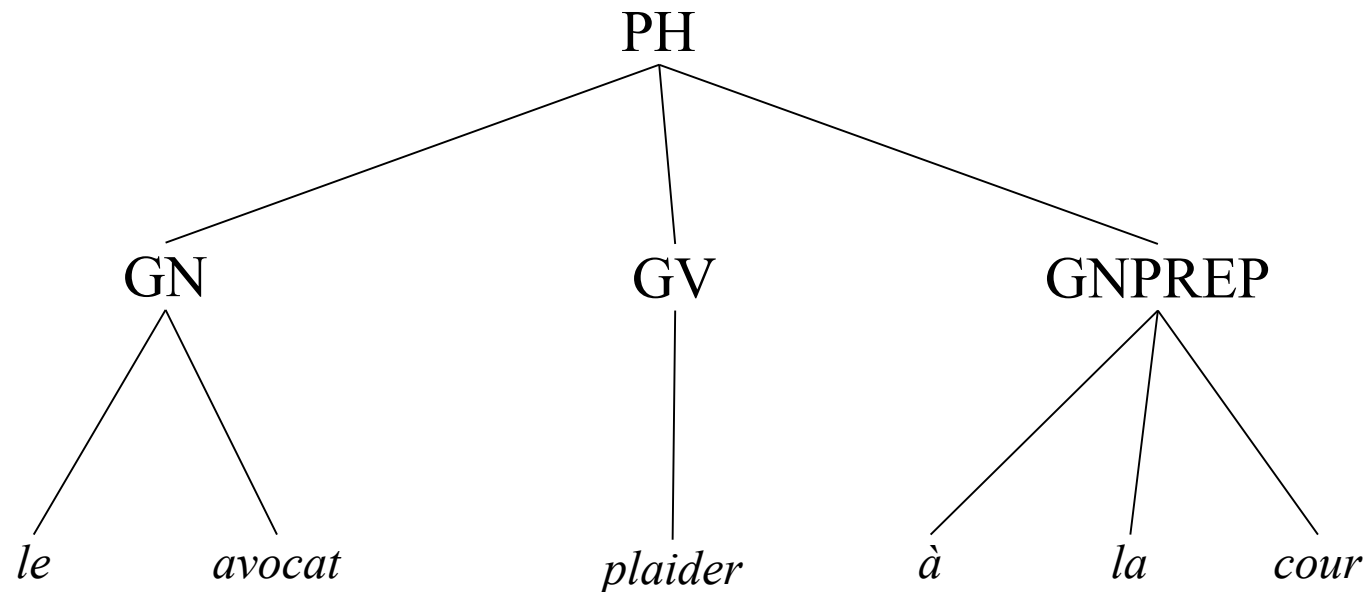


Arborescence

structure + jeu de variablesinstanciées

Application SYGFRAN (SYGMART) [chauché]

“L’avocat plaide à la cour”



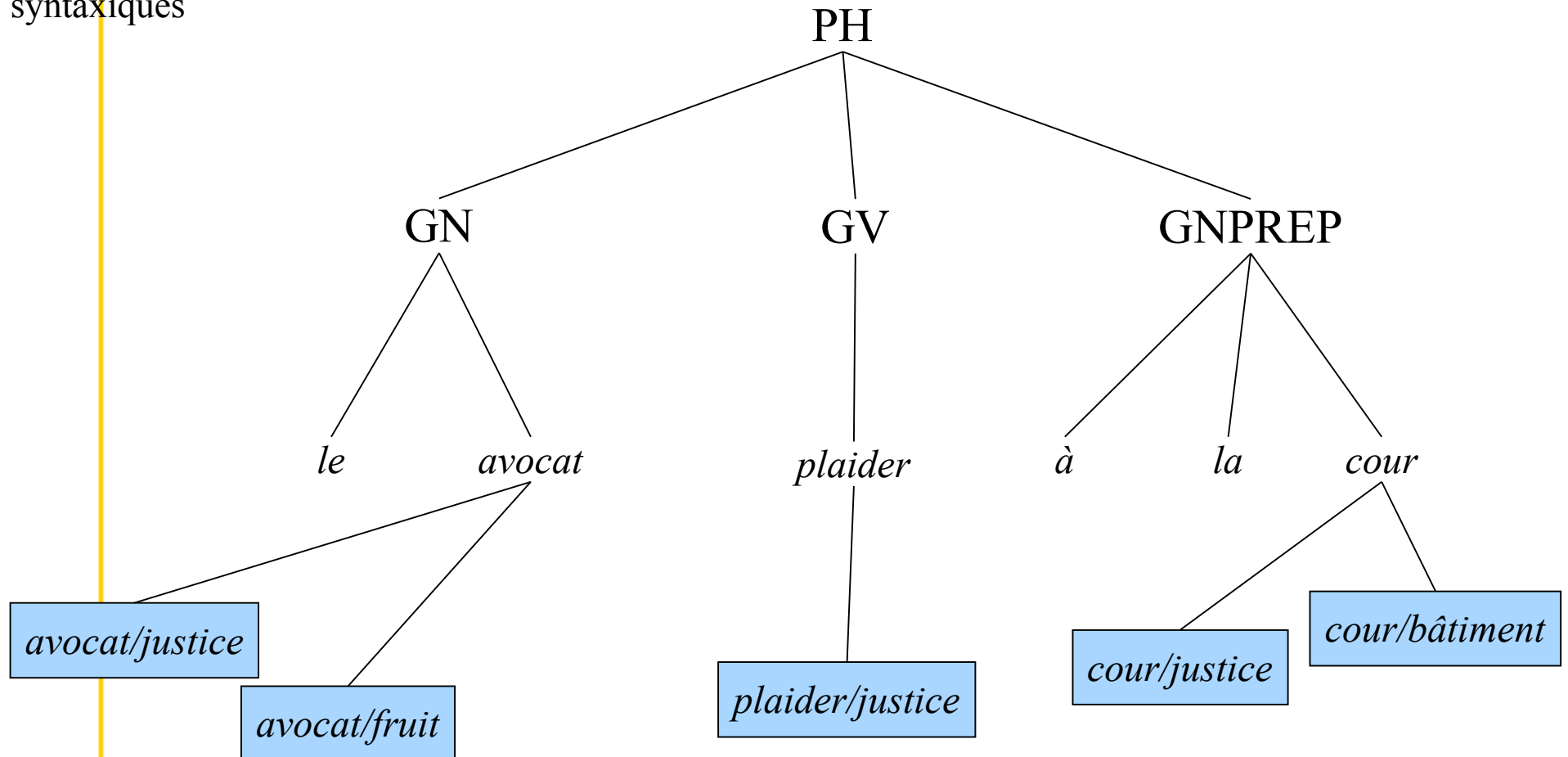


Transformations (1/3)



Arbres
morpho
syntaxiques

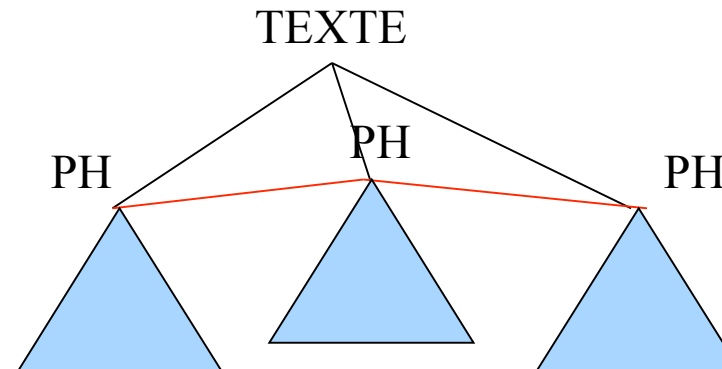
Ajout des acceptions



Arbres
morpho
syntaxiques

Transformation en graphe

Liens interphases



Calcul de distance entre phrases

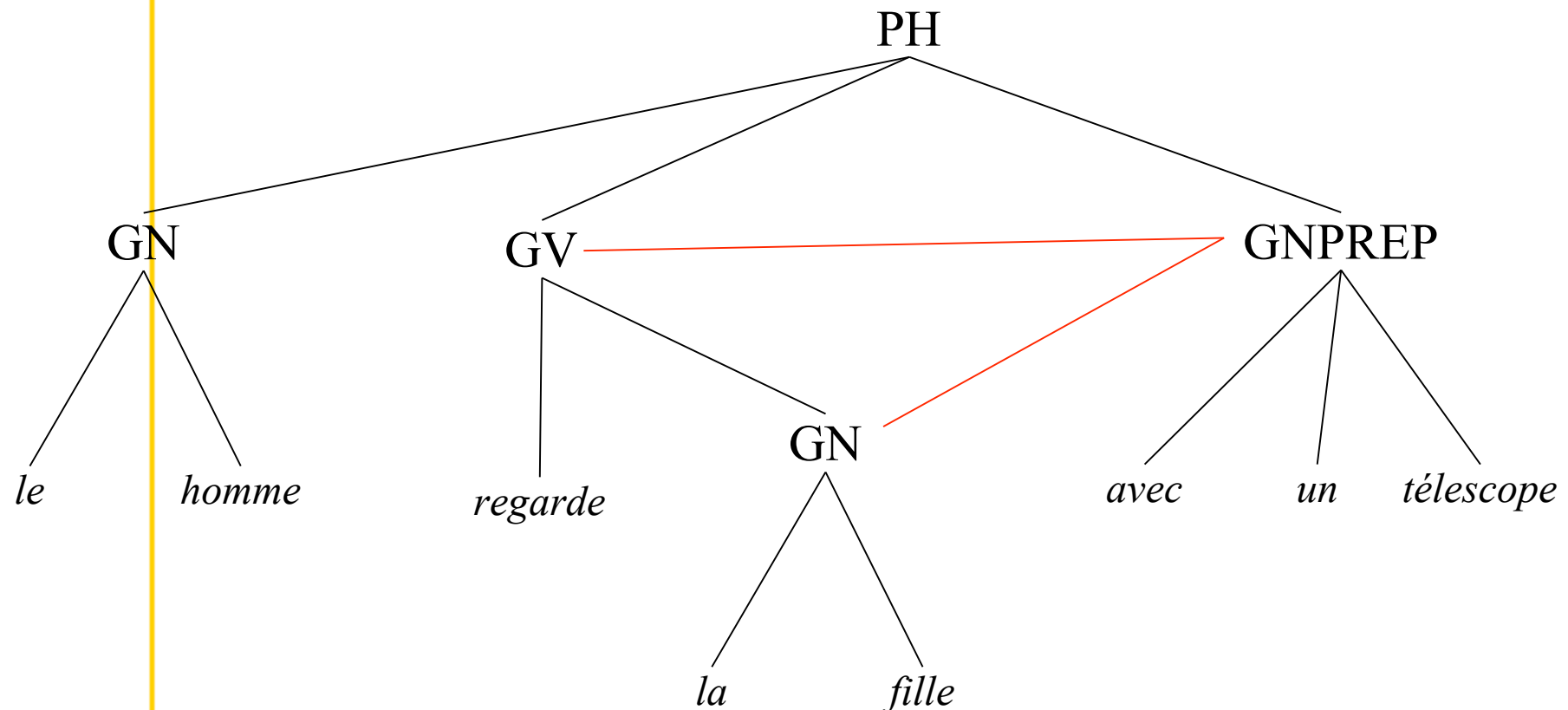
Ajout d'informations de surface aux informations de dépendances

Combinaison entre approches localistes et structurelles

Arbres
morpho
syntaxiques

Transformation en graphe

Énumération des rattachements possibles des GNPREP



Systeme Multi-Agents réactifs

TSP - recherche opérationnelle - réseau - bioinfo

[dorigo] [bertelle] [bonabeau, théraulaz] [bruten] [costa, hertz]

Principe

Stygmergie = communication indirecte
par modification de l'environnement

Dépot de phéromones - Evaporation lente

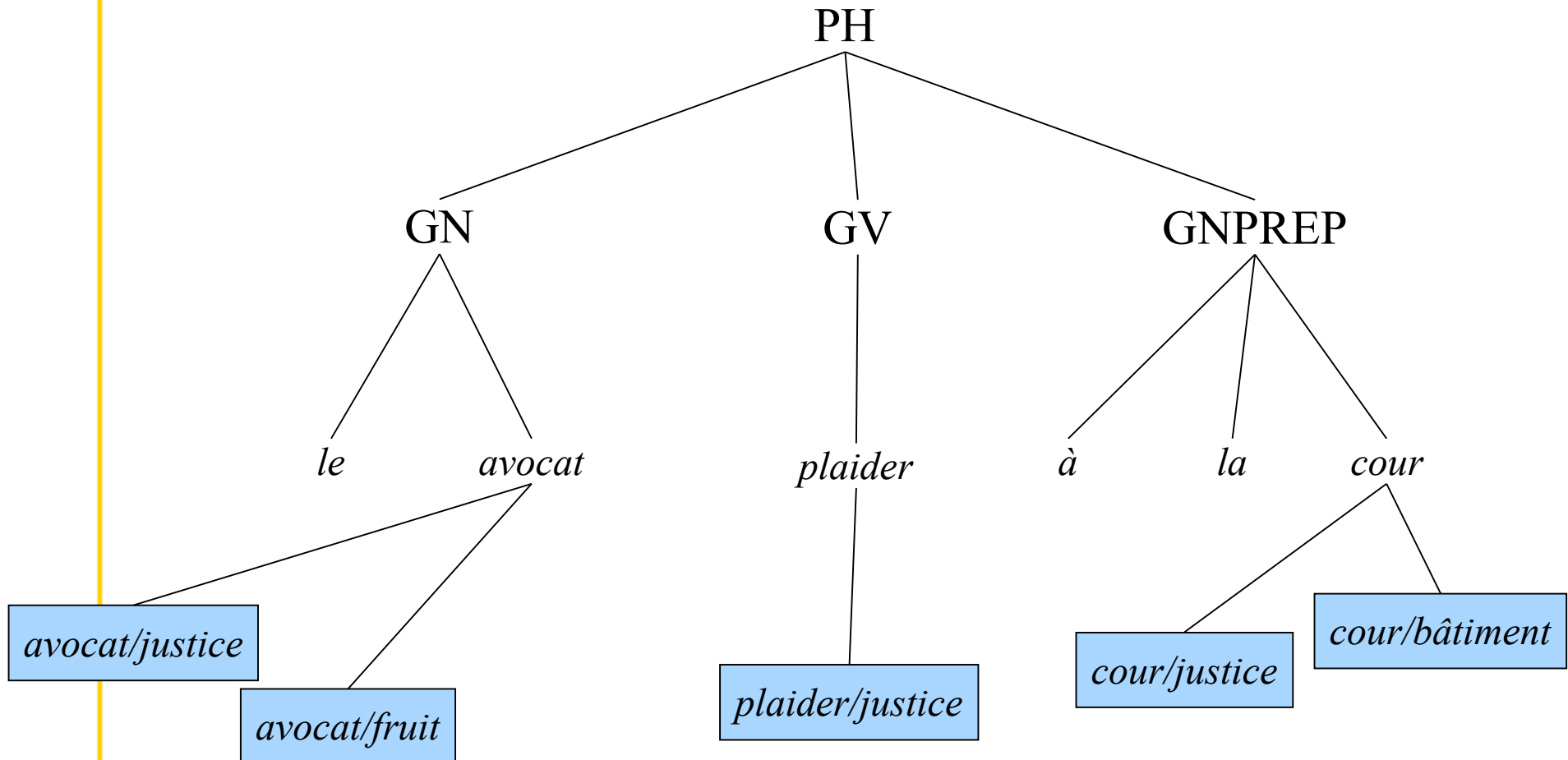
TALN

Agents cognitifs [stéphanini] [sabah]

sinon [lafourcade, guinand] [+cunningham] [zamora]

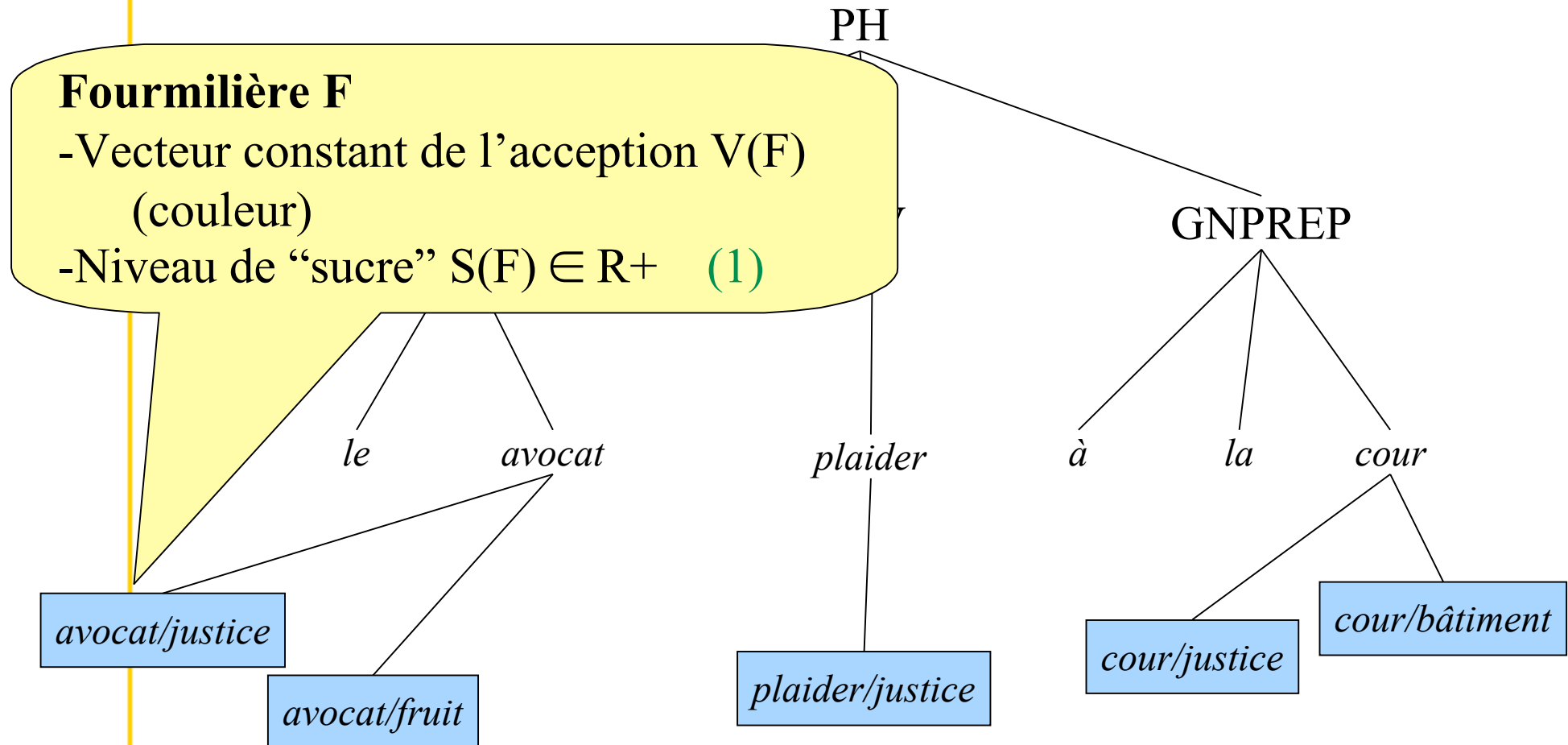


Algo
à fourmis





Algo
à fourmis





Algo
à fourmis

Noeud standard N

- Vecteur $V(N)$ (coloration) (V unitaire)
- Niveau de "sucre" $S(N) \in \mathbb{R}^+$ (1)

le *avocat*

avocat/justice

avocat/fruit

PH

GV

plaider

plaider/justice

GNPREP

à

la

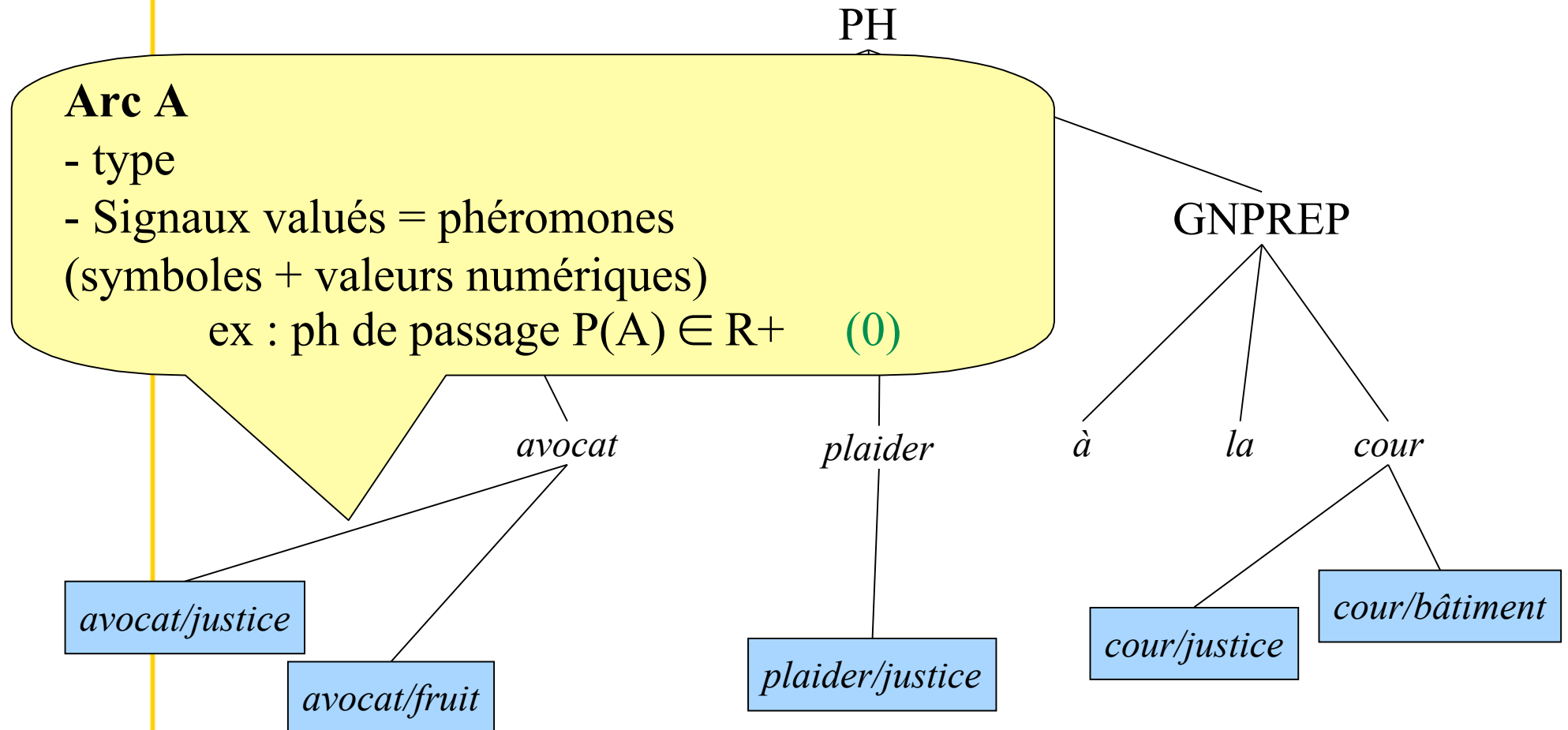
cour

cour/justice

cour/bâtiment



Algo
à fourmis

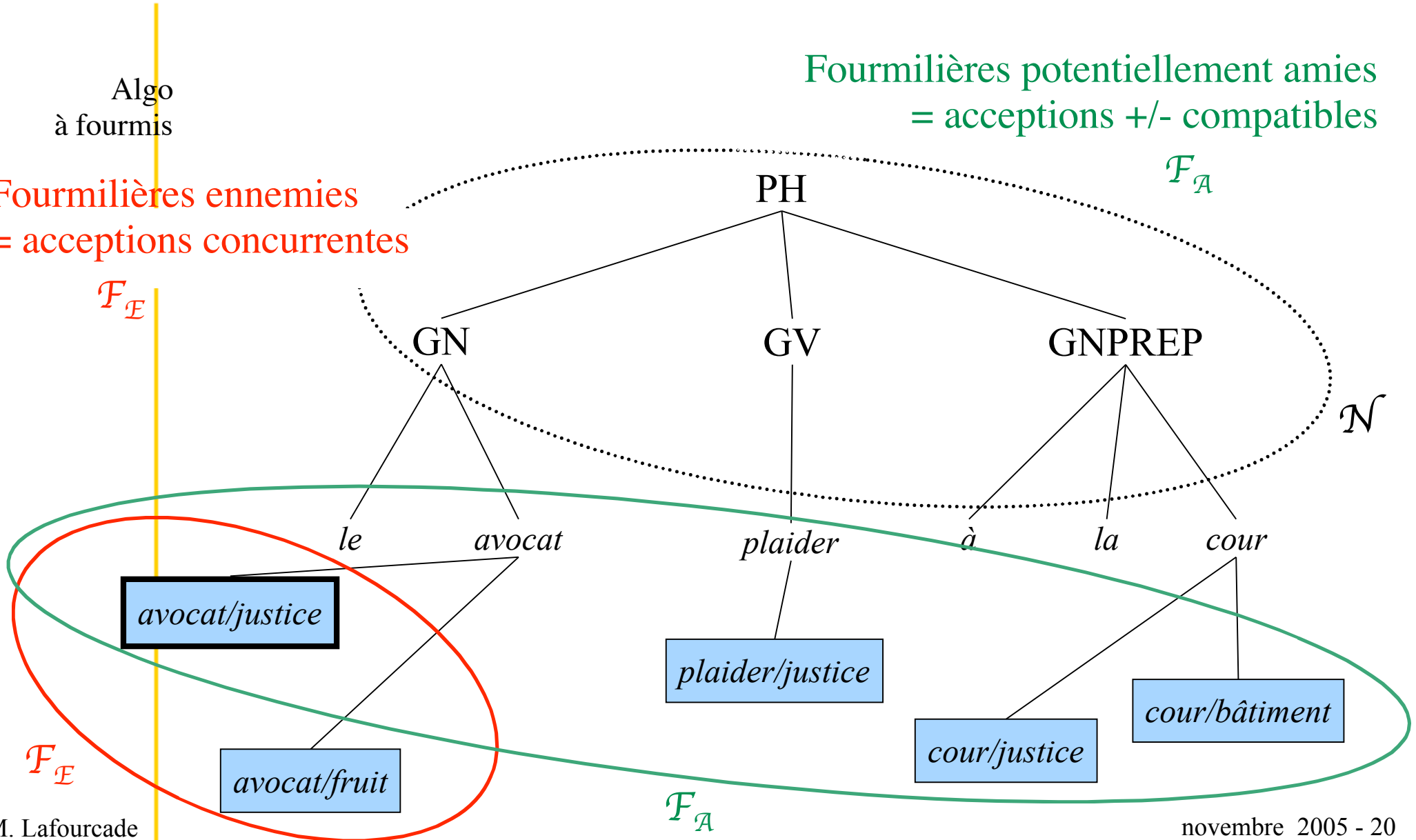




Algo
à fourmis

Fourmilières potentiellement amies
= acceptions +/- compatibles

Fourmilières ennemies
= acceptions concurrentes



Algo
à fourmis

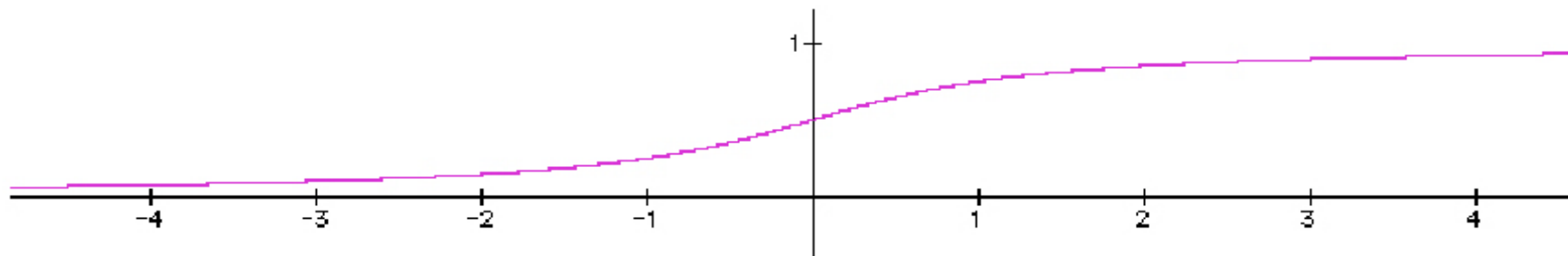
Production de fourmis

une fourmi à chaque cycle de la simulation

Selon une probabilité liée au niveau de “sucre” $S(F)$

fonction “sigmoïde”

coût de production (expérimentalement 1/10)



Algo
à fourmis



Attributs d'une fourmi f

Durée de vie : **expérimentalement 20 cycles**

Restitution de la ressource à la mort

Quantité de "sucre" transporté : **Q réel dans $[0,1]$**

Référence à la fourmilière d'origine F : **$V(f) = V(F)$**

Mode de déplacement courant : **$Prob = Q$**

Deux modes de déplacement

Recherche de "sucre"

Retour vers la fourmillère

Algo
à fourmis

Déplacement pseudo-aléatoire de $f \in F$

Evaluation des destinations possibles + liste tabou ($k = 1$)

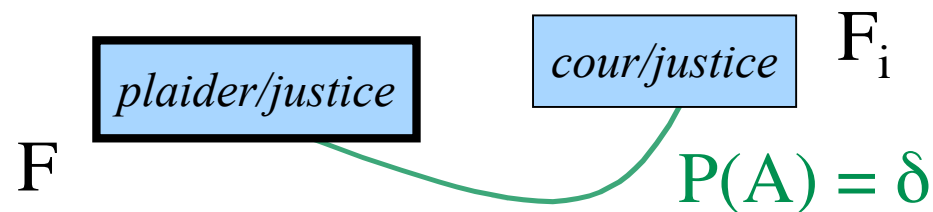
Modification de l'environnement

Dépot de phéromone sur l'arc traversé : $P(A) = P(A) + \delta$

Coloration du vecteur du nœud atteint : $V(N) = V(N) + \alpha V(f)$

somme normée et α et δ petits

Création d'un pont d'une fourmillière $F_i \in \mathcal{F}_A$ vers F



Algo
à fourmis



Principe général

Position courante = un nœud N avec k voisins N_i via des arc A_i

Fonction d'évaluation de chaque $N_i = Eval(N_i)$

Fonction d'évaluation de chaque $A_i = Eval(A_i)$

Fonction d'évaluation globale de la destination $dest_i$

$$Eval(dest_i) = \max \begin{cases} Eval(N_i) + Eval(A_i) \\ \epsilon \end{cases}$$

$$Prob(dest_i) = \frac{Eval(dest_i)}{\sum_{k=1}^n Eval(dest_k)}$$

Algo
à fournis

Mode recherche

$$Eval(N_i) = S(N_i)$$

$$Eval(A_i) = 1 - P(A_i)$$

action spécifique : prendre du sucre sur $N_i \in \mathcal{N} \cup \mathcal{F}_E$

Mode retour

$$Eval(N_i) = 1 - \frac{2}{\pi} D_A(V(N_i), V(f))$$

$$Eval(A_i) = P(A_i)$$

action spécifique : déposer du sucre sur N_i si $N_i \in \mathcal{F}_A$





Algo
à fourmis



A chaque cycle les phéromones $P(A)$ diminue
par “évaporation”

Arc entre deux noeuds de l’arbre

$$P(A) = \text{Min} (P(A) - \eta, 0)$$

Pont en N_a et N_b

$$P(A) = \text{Min} (P(A) - d * \eta, 0)$$

$$d = \text{distance_ultramétrique}(N_a, N_b)$$

Le pont disparaît si $P(A) = 0$

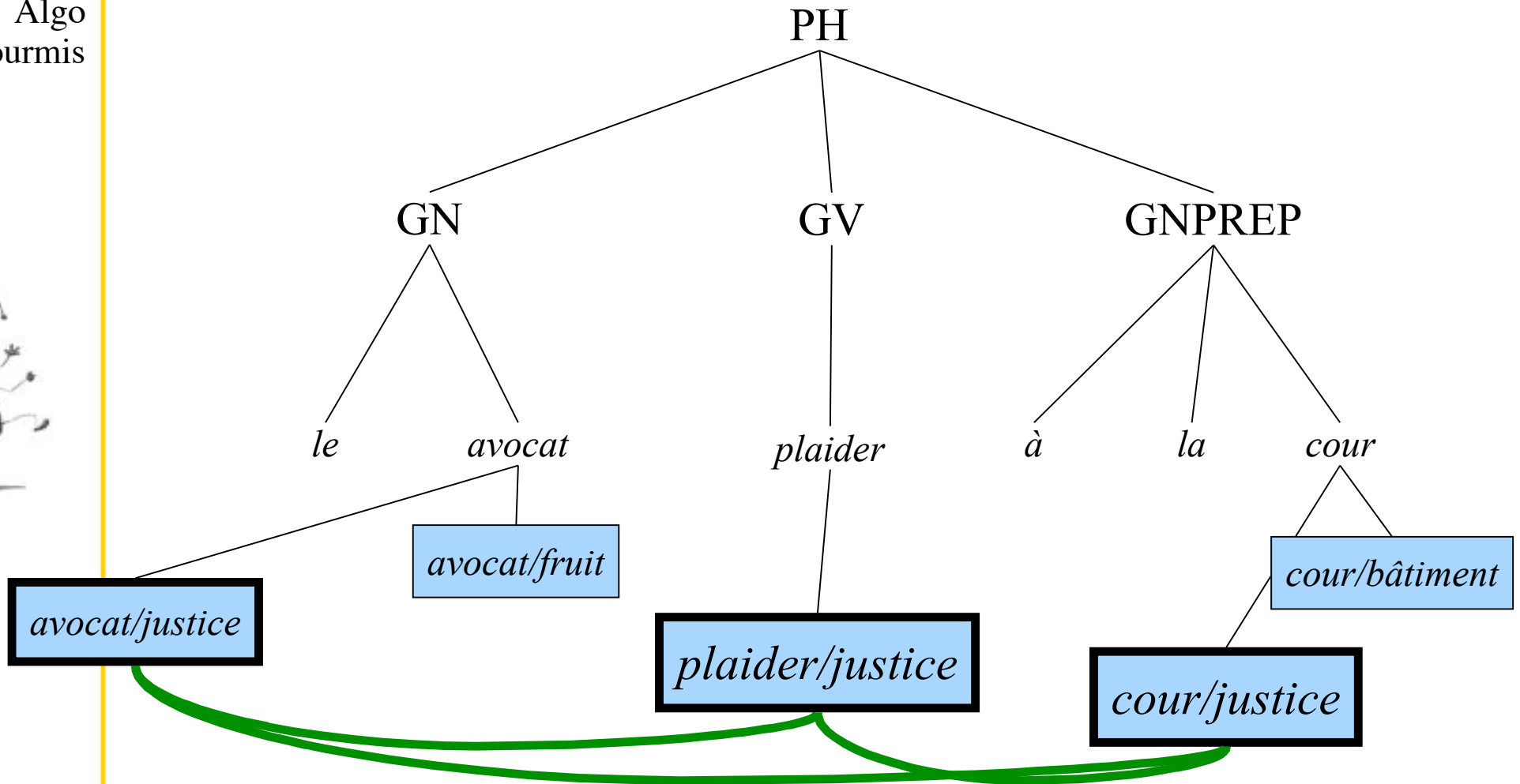
Plus un pont est long, plus il est difficile à maintenir !

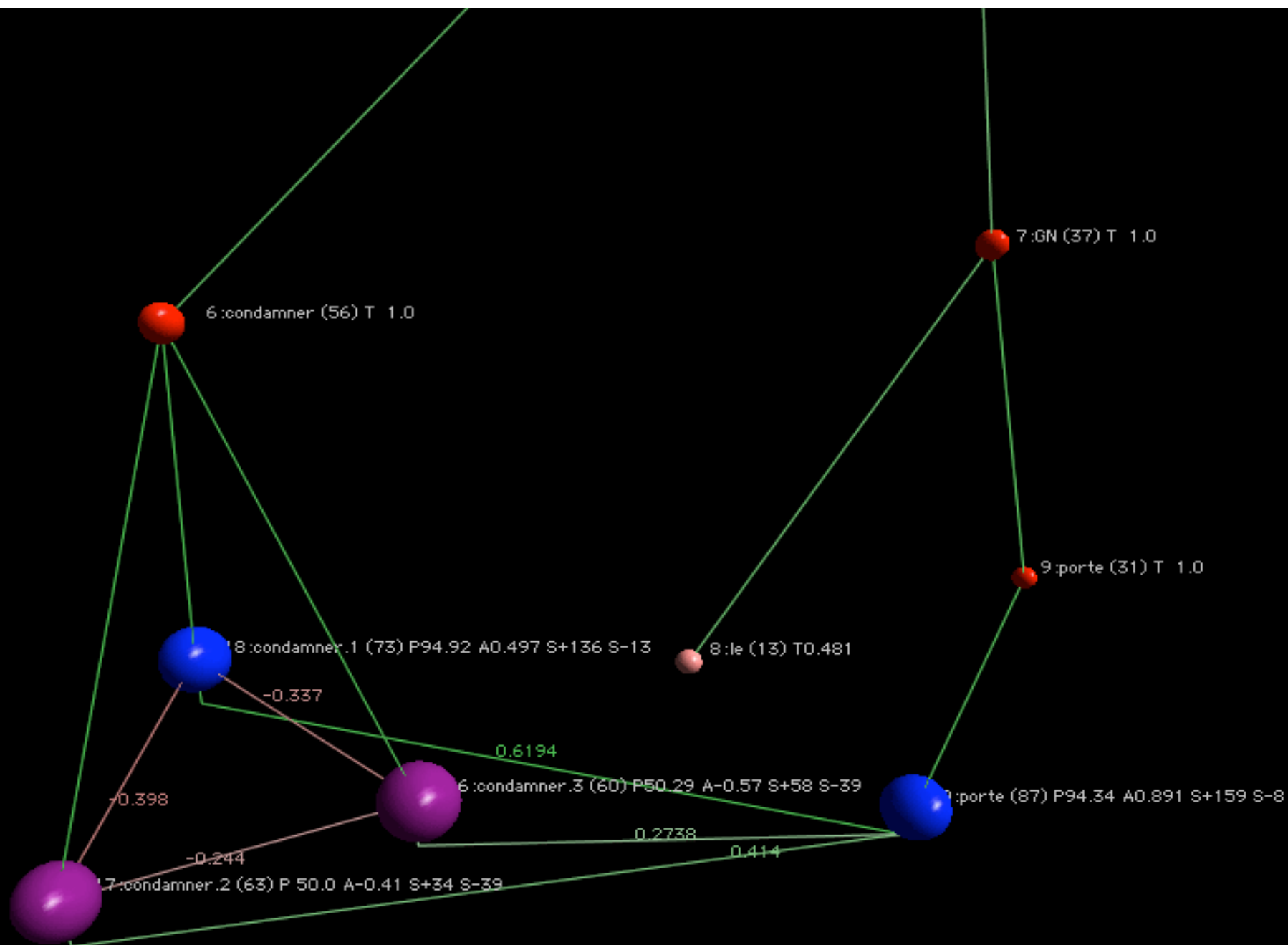


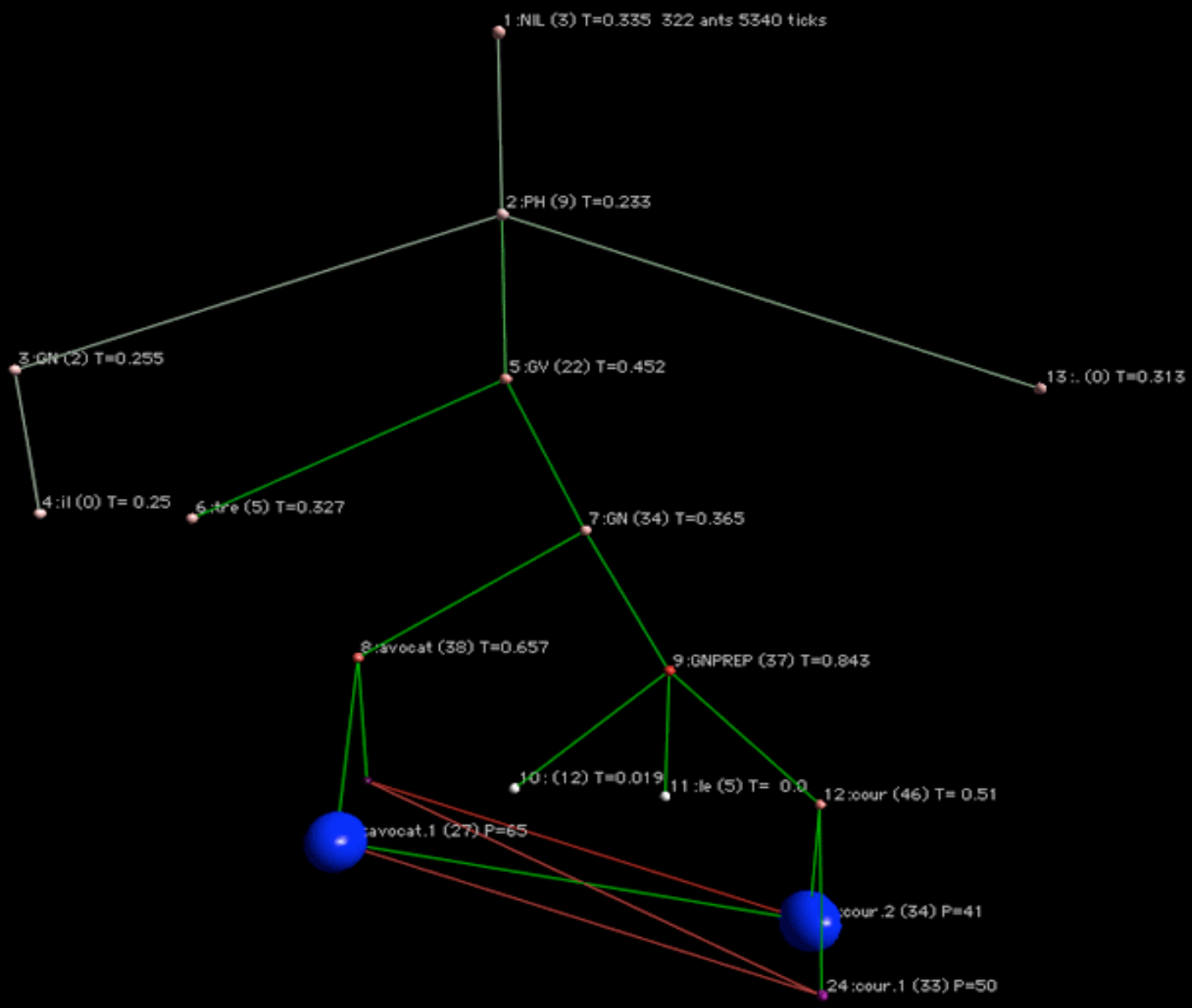
Exemple

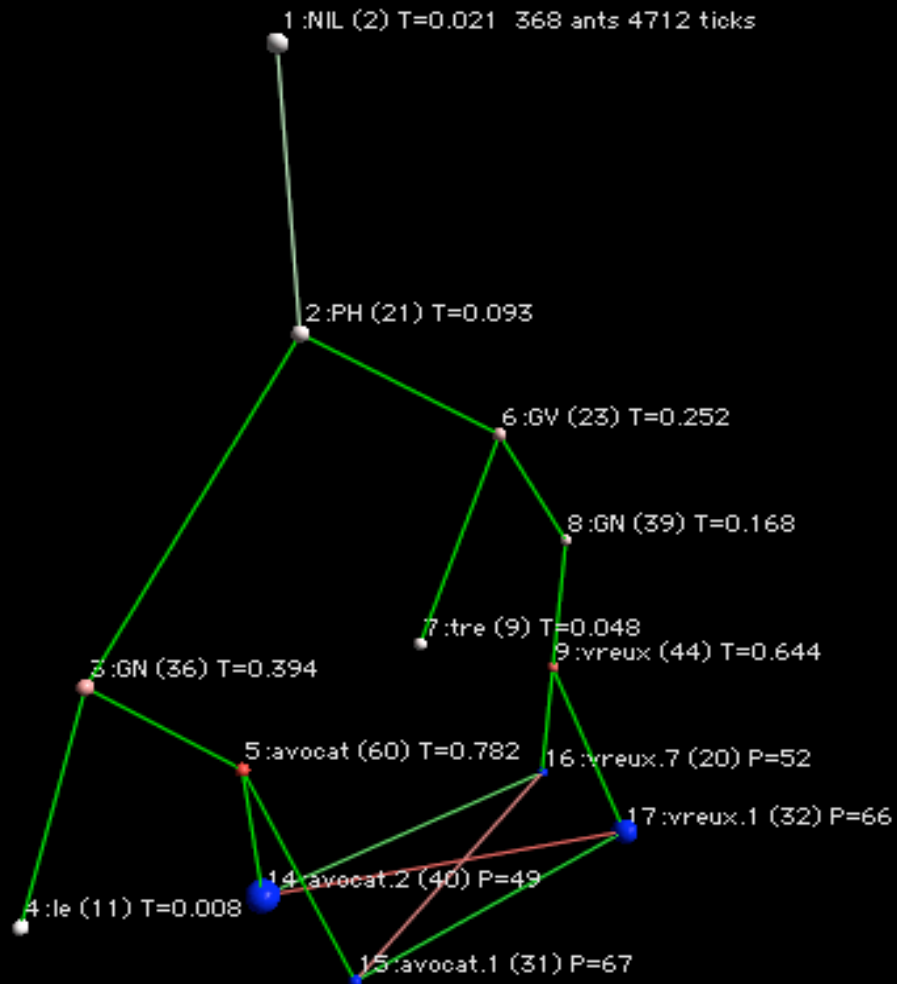


Algo
à fourmis





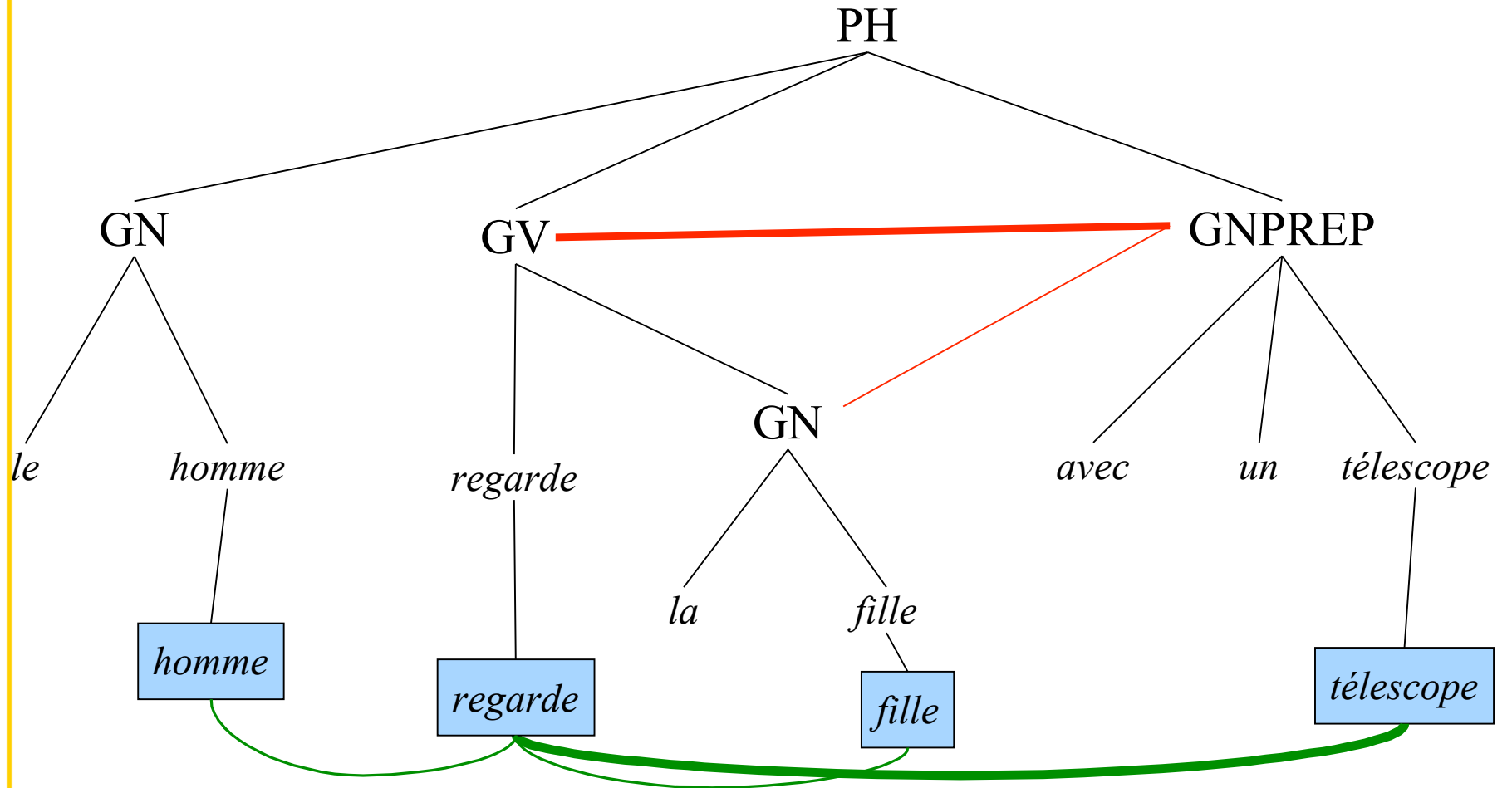






Exemple de rattachement

Algo
à fourmis





LIRMM



Centre National de la
Recherche Scientifique



Université
Montpellier II

Algo
à fourmis



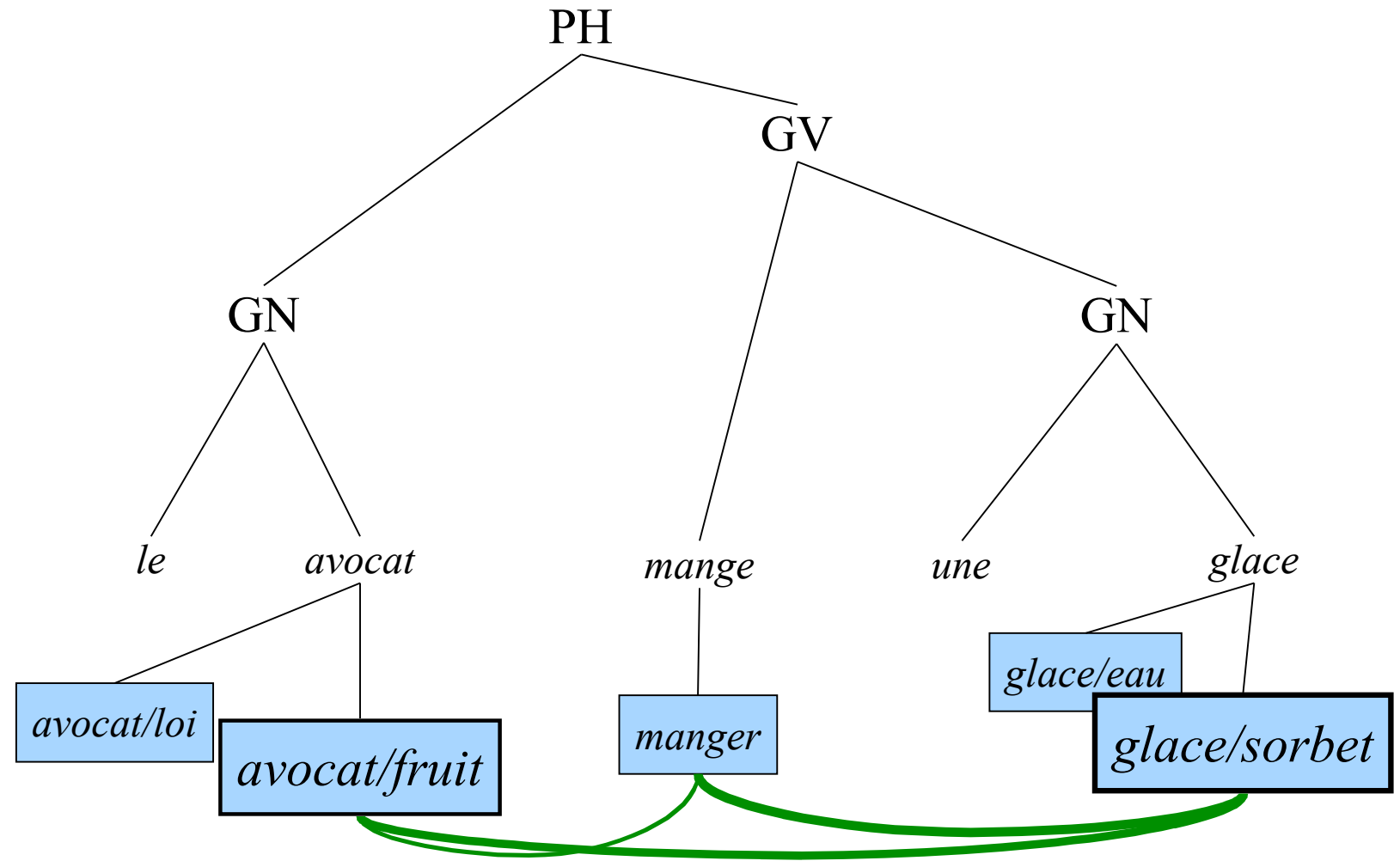
Et ça marche tout le temps ?



Ca marche pas !

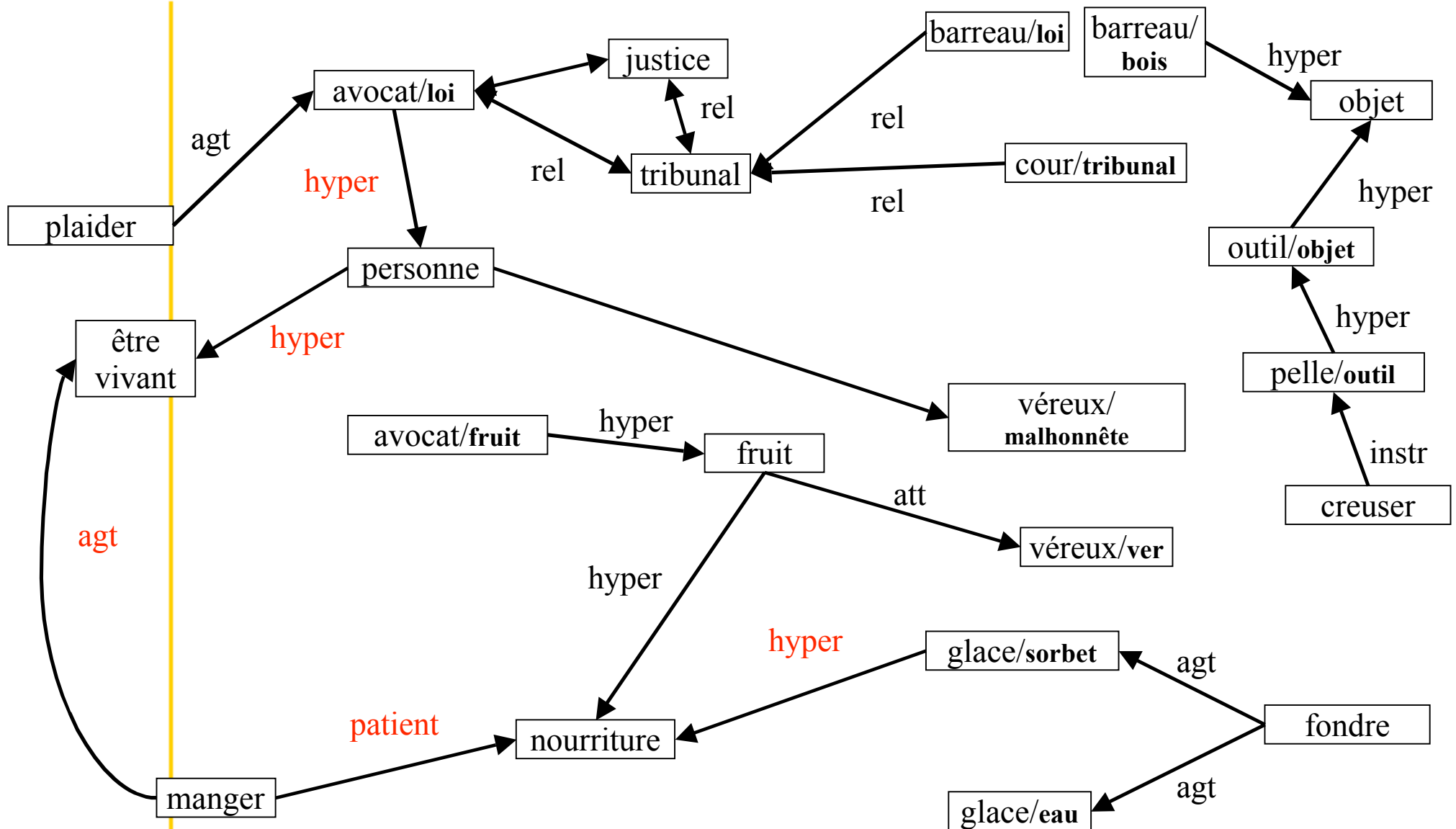


Algo à fournis



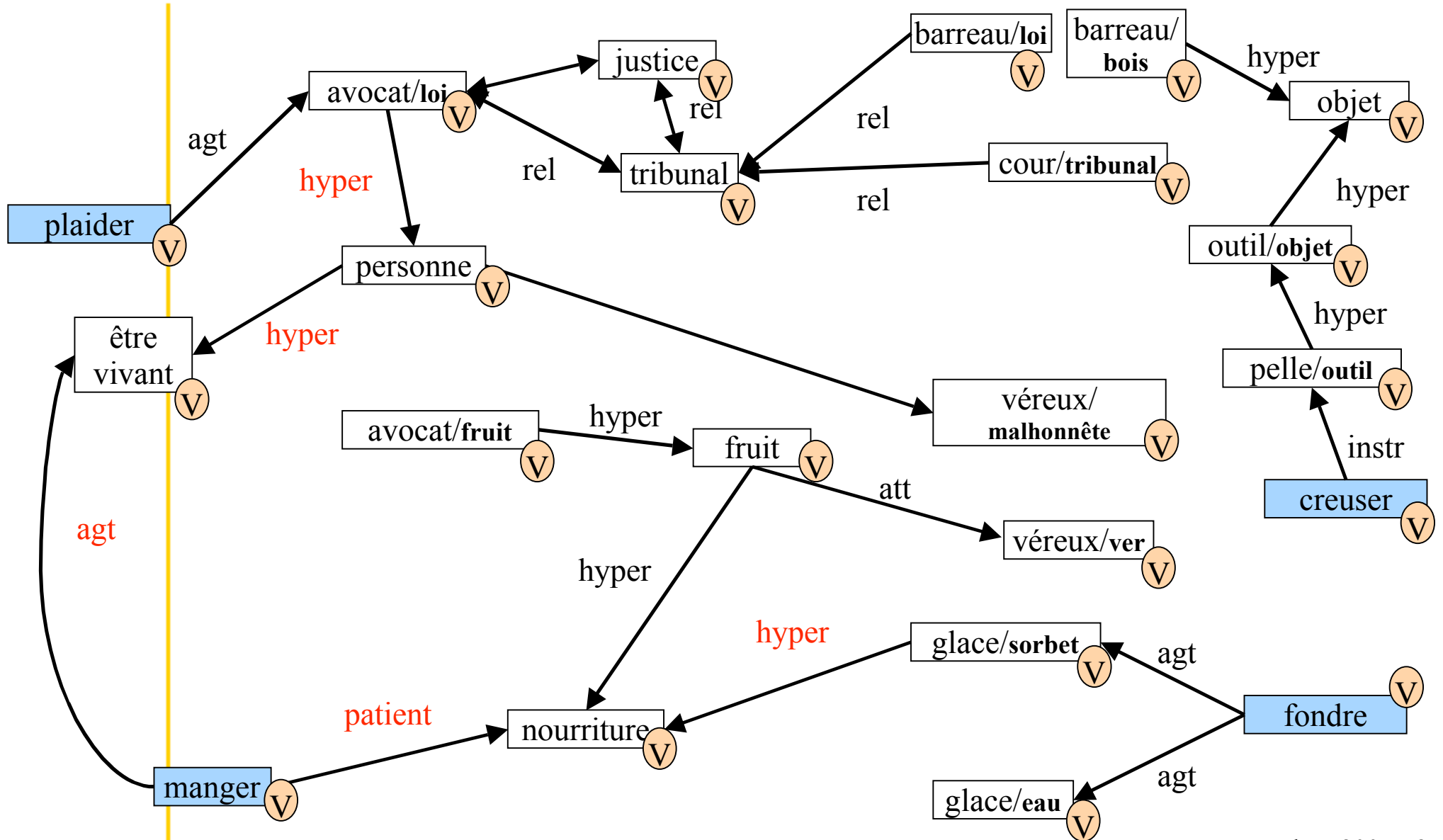


Réseau lexical





Réseau lexical



Castes

Fourmis avec fonction d'évaluation particulière

Castes exploitant le réseau lexical

Prédicat recherche agent

Prédicat recherche patient

Prédicat recherche instrument ...

Recopie locale

Acceptions et relations du réseau

(désespérement)

Algo
à fourmis

Fourmis créées par un verbe (prédicat)

Mode recherche

$$Eval(N_i) = S(N_i)$$

$$Eval(A_i) = 1 - P(A_i) + \max \begin{cases} 2 & \text{si } type(A_i) = patient \\ 0 & \text{sinon} \end{cases}$$



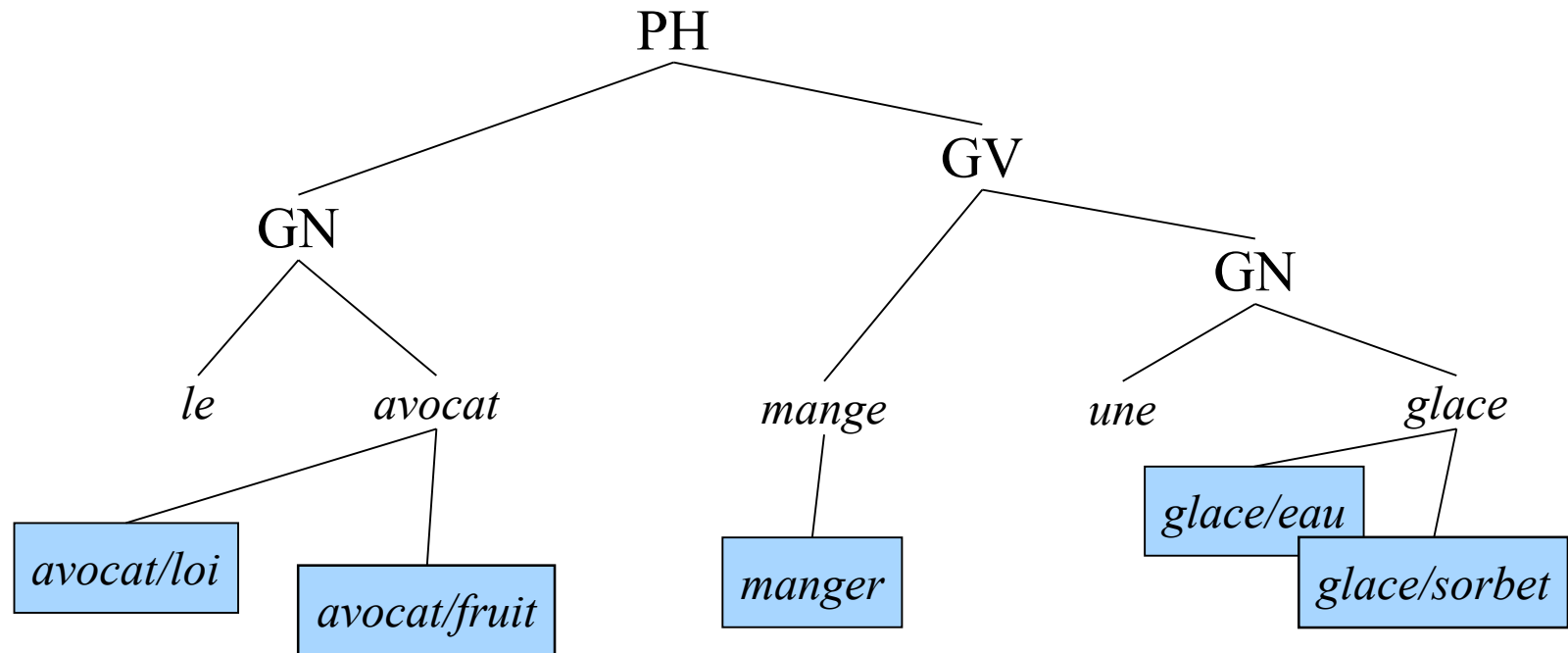
Mode retour

$$Eval(N_i) = 1 - \frac{2}{\pi} D_A(V(N_i), V(f))$$

$$Eval(A_i) = P(A_i) + \max \begin{cases} 2 & \text{si } type(A_i) = patient \\ 0 & \text{sinon} \end{cases}$$

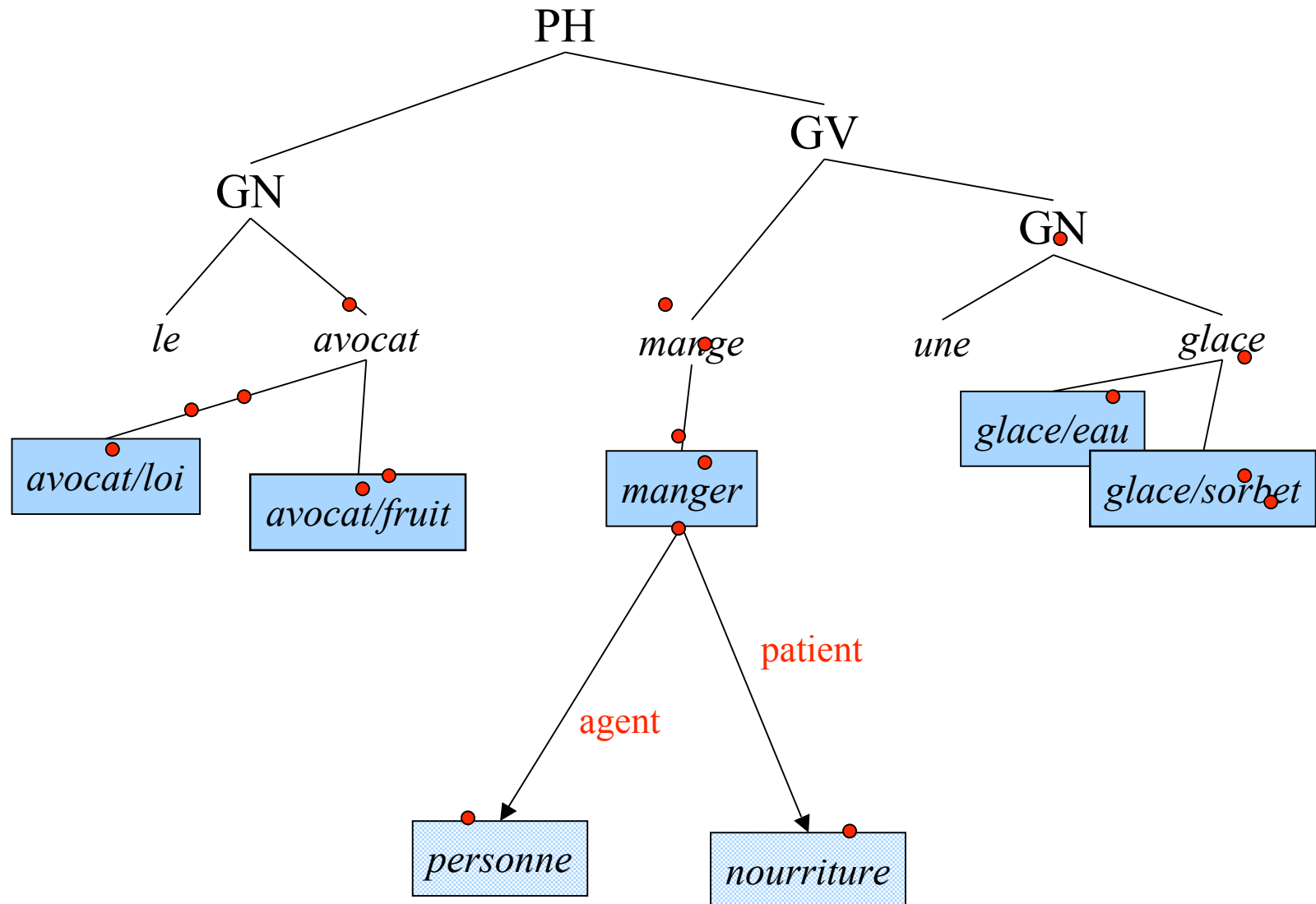


Algo
à fourmis



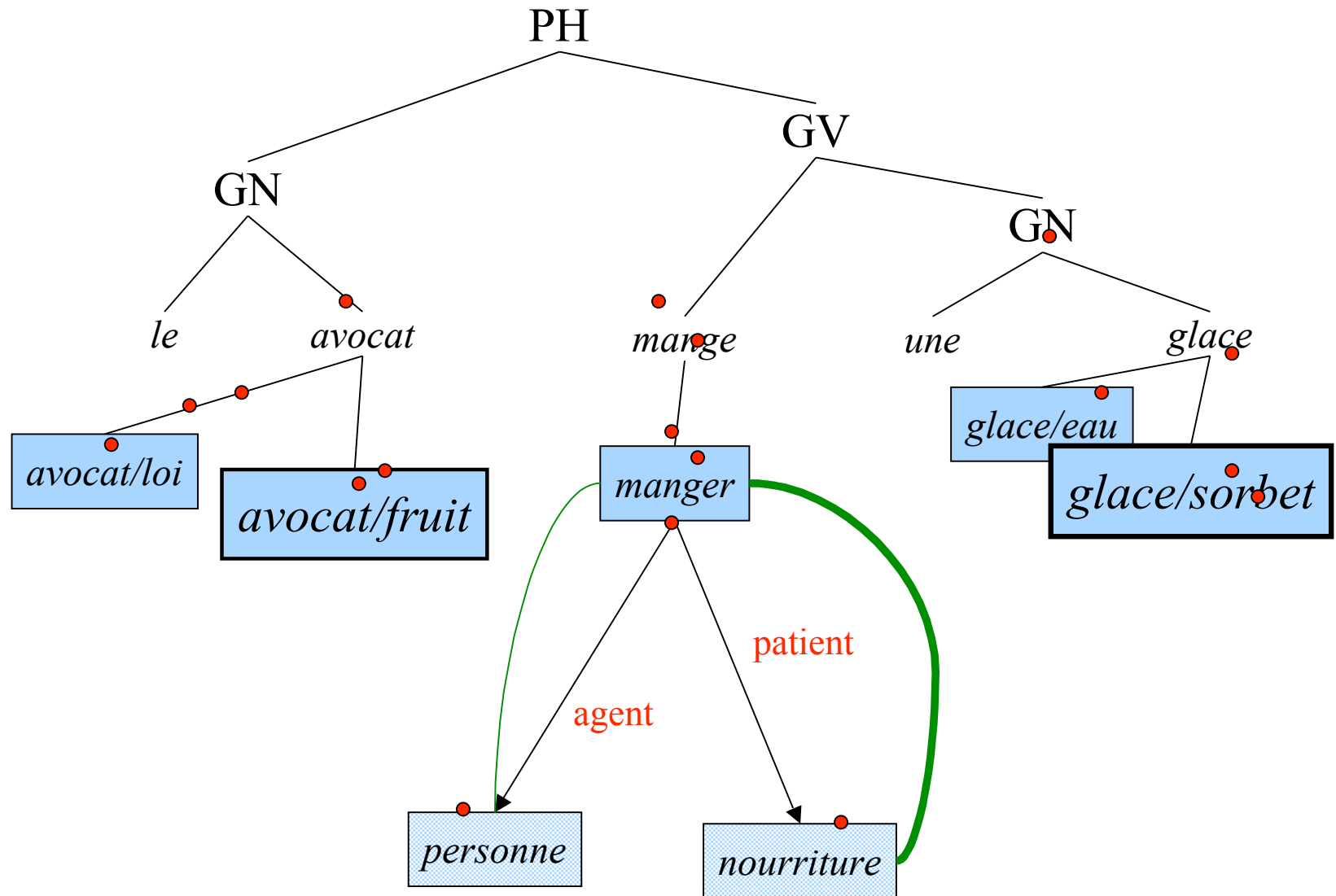


Algo
à fourmis



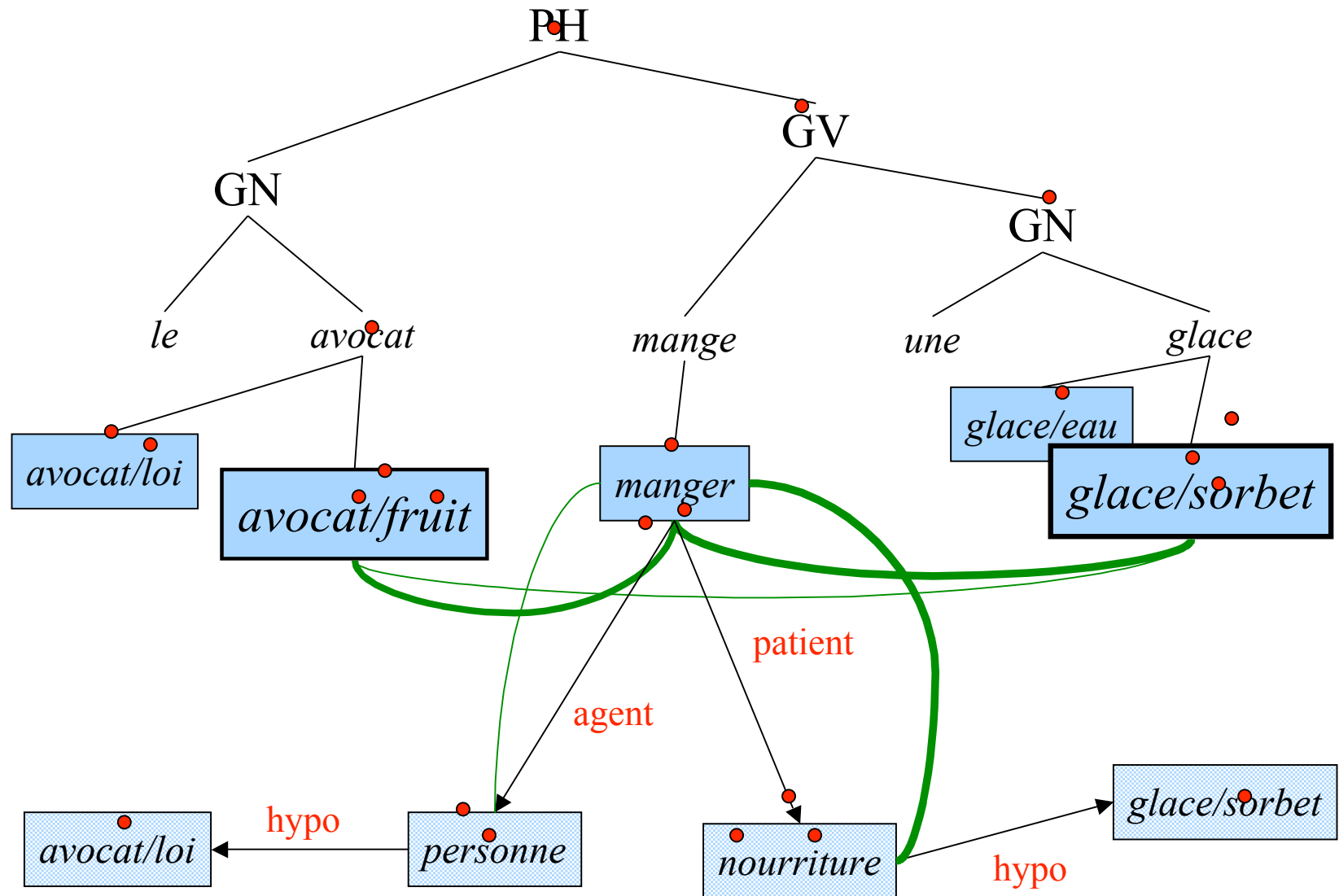


Algo
à fourmis



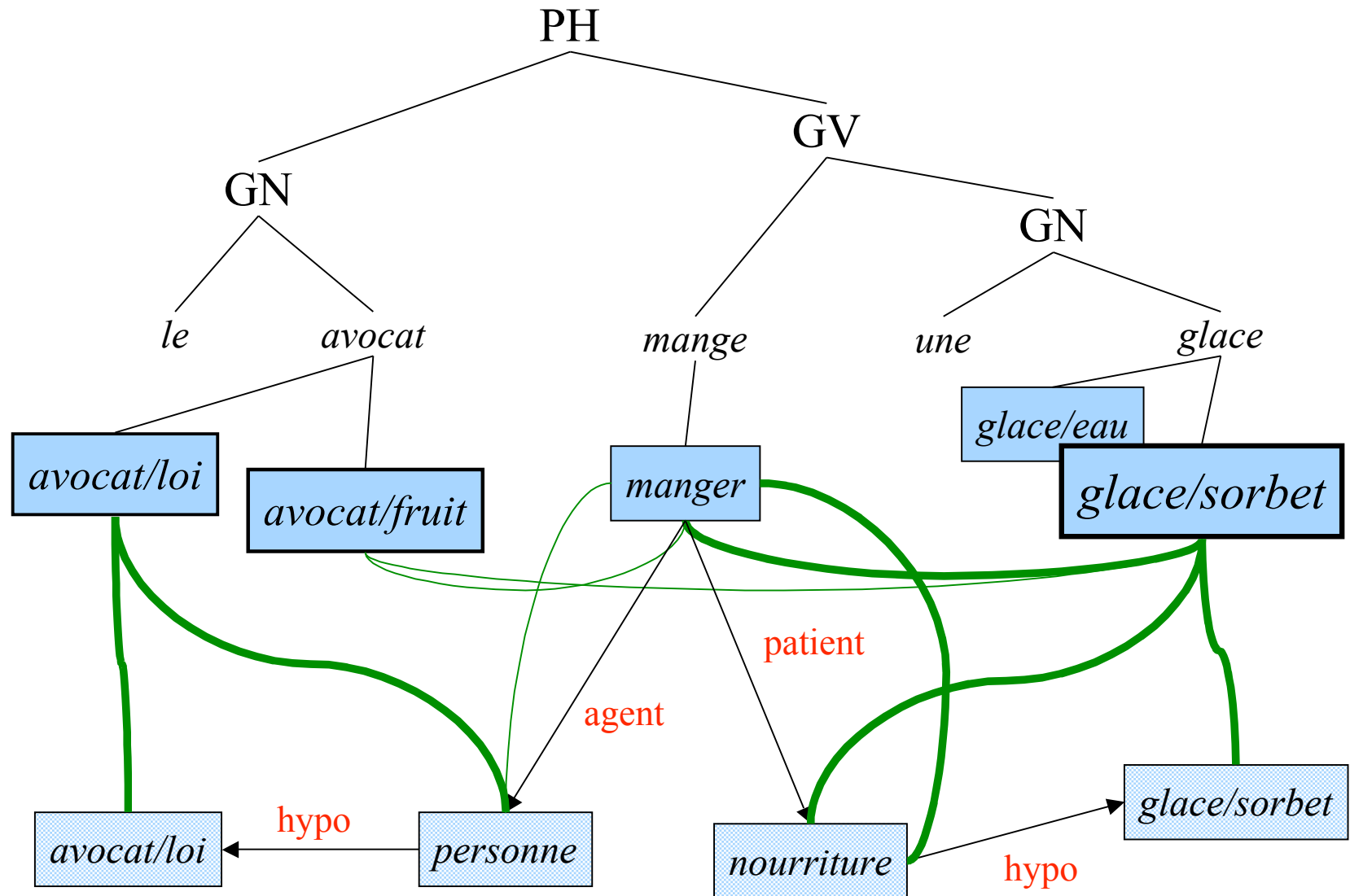


Algo
à fournir



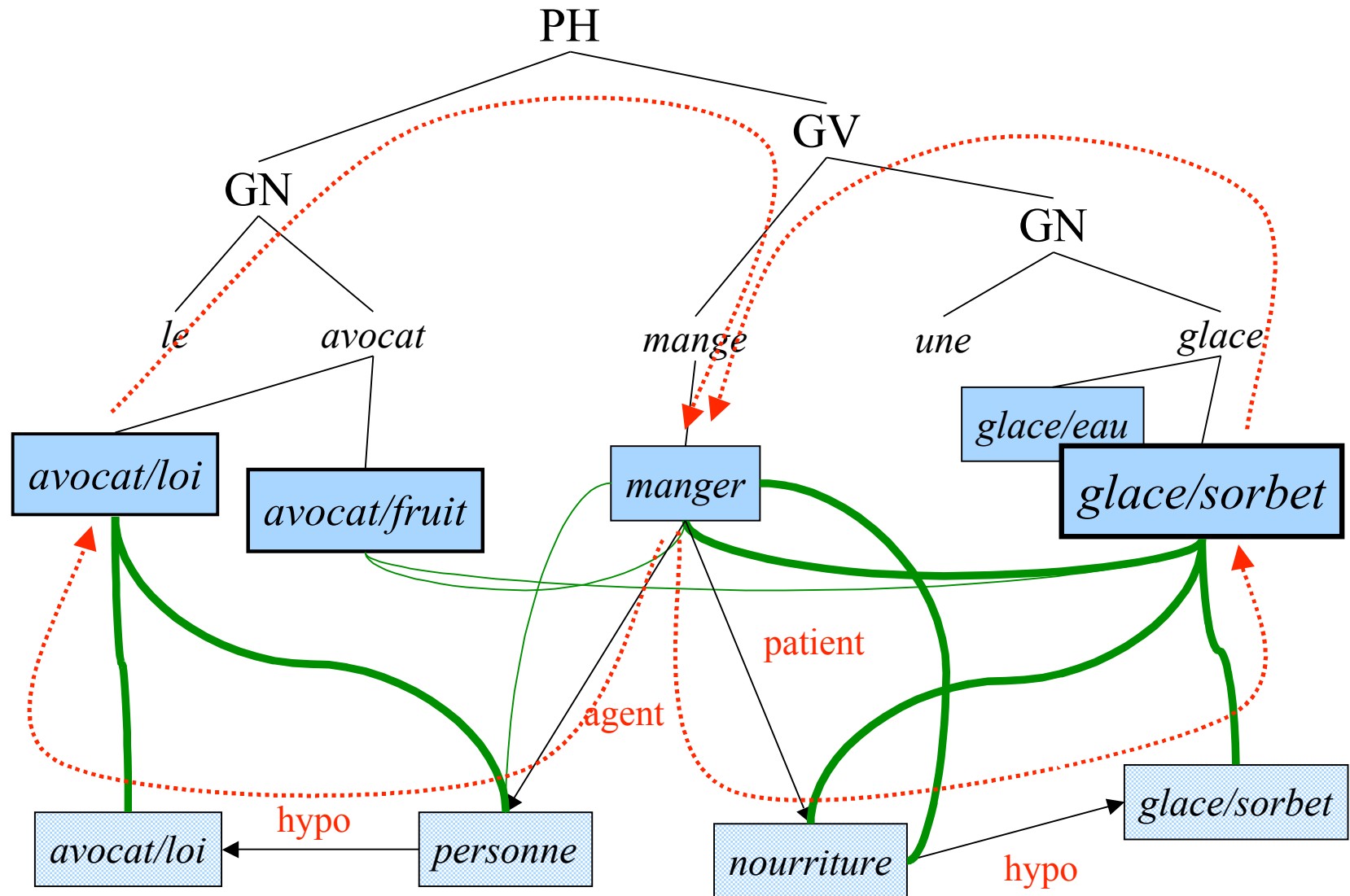


Algo
à fourmis



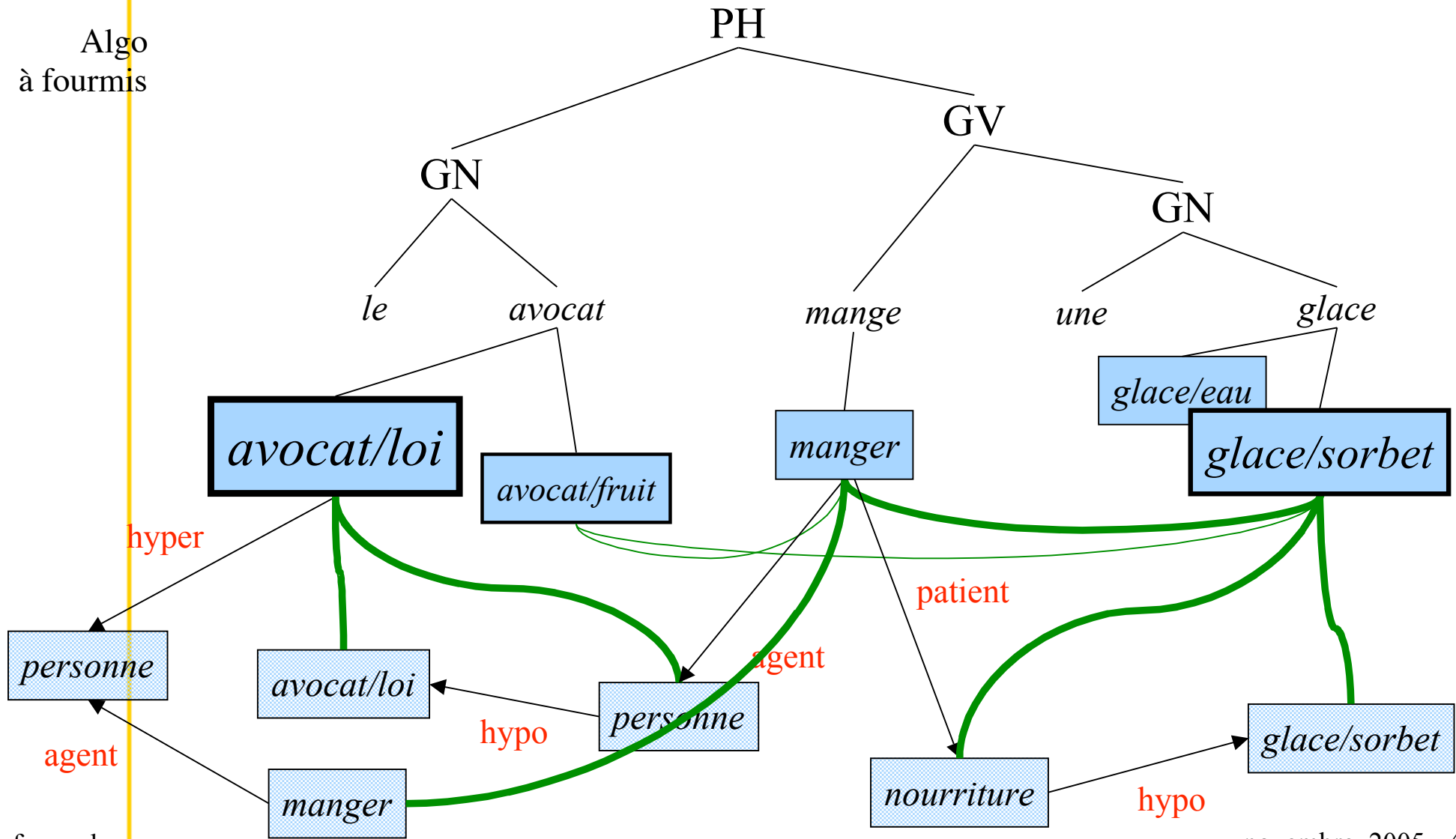


Algo
à fourmis





Algo
à fourmis

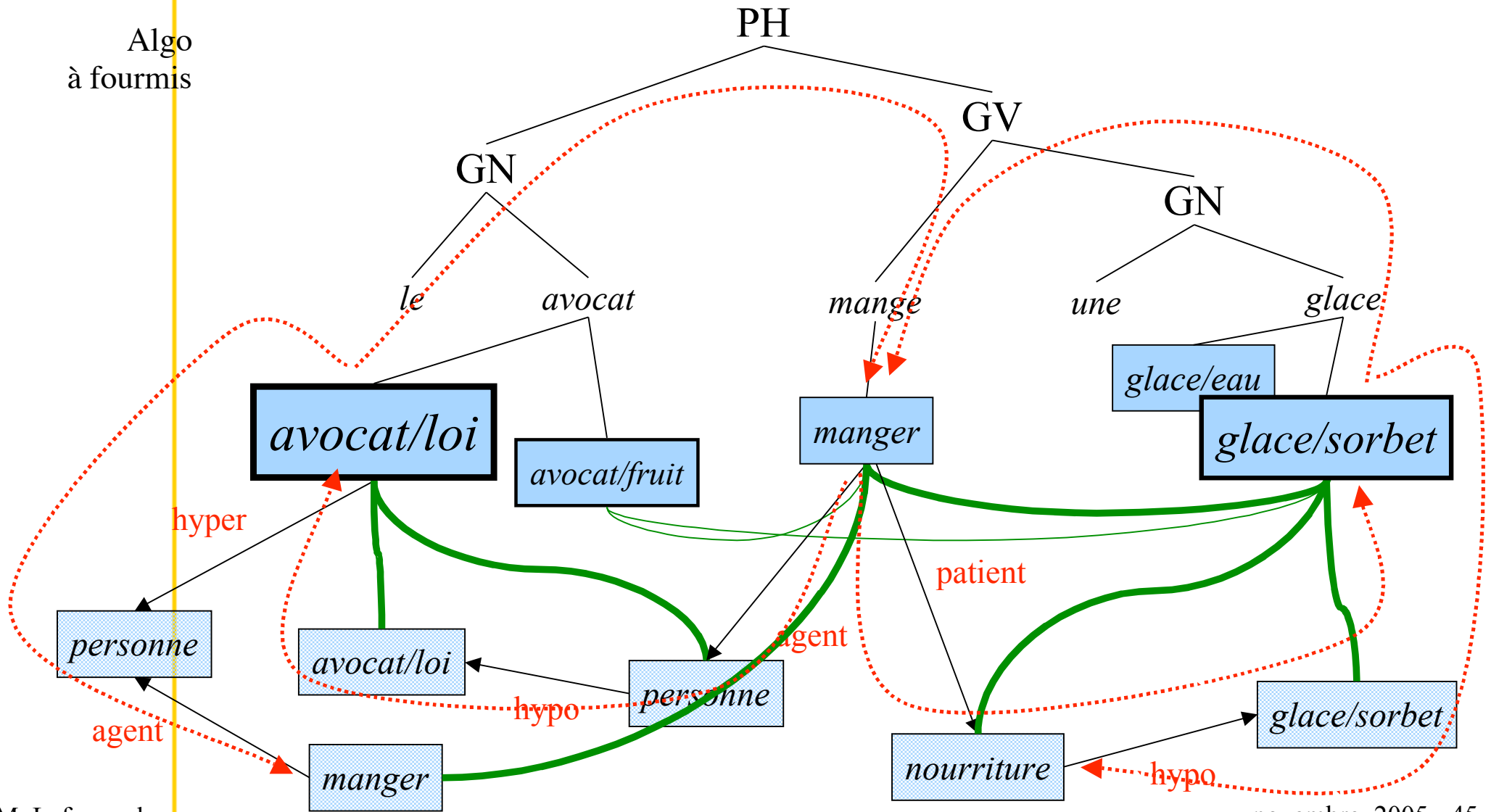




Ca marche !



Algo
à fourmis

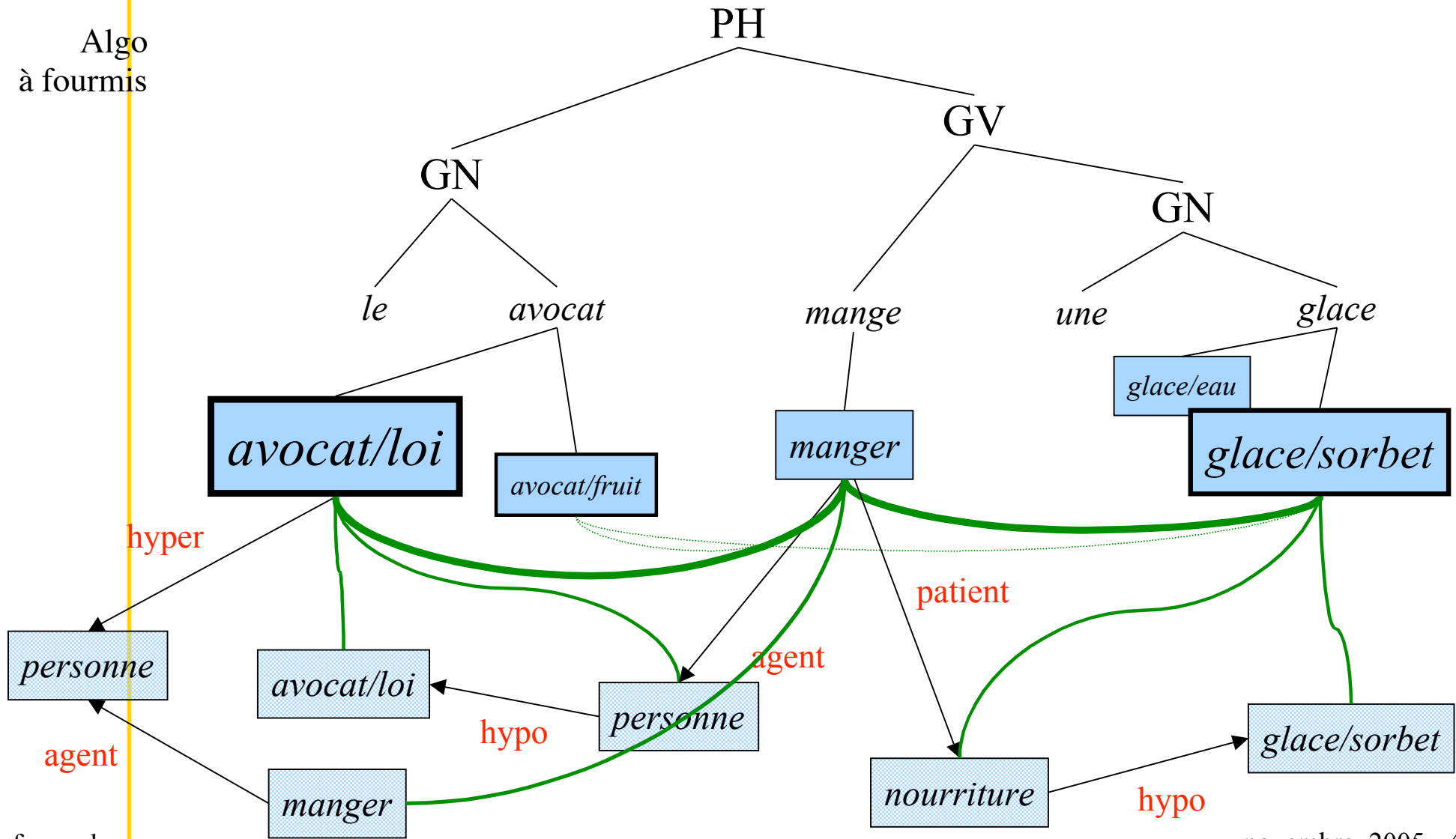




Ca marche !



Algo
à fourmis





Taux de production d'une caste donnée

Ajustement par récompense

Auto-arrêt du système

Par facteurs interne (\neq recuit simulé)

Fourmis « gelantes » - réduit le taux d'évaporation des phéromones

Produites si environnement stable
vu depuis la fourmilière

Inhibition

Fourmis tueuses - phéromone d'alerte



Inclusion de schémas fréquents

Rattachement prépositionnel [gala, lafourcade]

Distribution d'acceptions

en général

sur un domaine ?

Fréquence de noms composés

petit gris - serpent de mer

Fréquence pour les locutions

mettre x sur la paille



Emergence d'une solution par exploitation d'erreurs

Problème de coordination facilité

Architecture générique et extensible

Difficulté à paramétrer le système